

DOCUMENT RESUME

ED 352 419

TM 019 355

AUTHOR Kleinsasser, Audrey M.; Horsch, Elizabeth A.
TITLE Teaching, Assessment, and Learning: Invitation to a Discussion.
INSTITUTION Wyoming Univ., Laramie. Wyoming Center for Educational Research.
SPONS AGENCY Wyoming State Dept. of Education, Cheyenne.
PUB DATE Aug 92
NOTE 167p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC07 Plus Postage.
DESCRIPTORS *Educational Assessment; Educational Research; Elementary Secondary Education; Evaluation Methods; *Evaluation Utilization; *Learning; Outcomes of Education; Policy Formation; *Student Evaluation; Teacher Role; *Teaching Methods; Test Use
IDENTIFIERS Alternative Assessment; *Outcome Based Education; Performance Based Evaluation; *Wyoming

ABSTRACT

Forms of assessment that are consistent with outcomes based education are discussed. Assessment in service to learning must inform students about their learning, provide data for teachers who plan instruction, inform parents about student progress, and provide data for policymakers. Chapter 1 of this document is "The Implications of Prevailing Testing Practices." Chapter 2, "Connecting Learning and Assessment," summarizes research data about learning. Chapter 3, "Grades, Tests, and Ordeals: The American Testing Culture"; and Chapter 4, "Roles, Rules, and Rituals in a Culture of Assessment: Blurring the Distinctions," show that practices that exist in a culture of testing will probably need to be recast in a culture of assessment. Chapter 5, "Making the Transition: The Assessment of Educational Standards and Outcomes," provides standards for outcomes-based education and examples of performance based assessment from teachers in Wyoming. A resource section contains 13 examples of strategies that teachers in Wyoming use to meet instructional outcomes through assessment, including video assessment; product testing; first grade multidisciplinary assessment; alternative assessment; monographs; a Renaissance fair; self-evaluated gymnastics; integration of computers; literacy assessment in third grade; reading and writing in first grade; assessing the integration of reading, writing, social studies, and science. Three summary sheets and 56 sources are included. (SLD)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

Teaching, Assessment, and Learning: Invitation to a Discussion

by

Audrey M. Kleinsasser

and

Elizabeth A. Horsch

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.



"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

VICKI L. ATTERBERRY

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

University of Wyoming
Wyoming Center for Educational Research

Teaching, Assessment, and Learning:

Invitation to a Discussion

by

Audrey M. Kleinsasser

and

Elizabeth A. Horsch

Published by

Wyoming Center for Educational Research

Box 3374

College of Education

University of Wyoming

Laramie, Wyoming 82071-3374

August, 1992

Portions of this document were developed as a result of activities conducted under a FY 1989 Chapter 2 ESEA contract between the University of Wyoming College of Education and the Wyoming Department of Education.

The University of Wyoming and the Wyoming Department of Education do not discriminate on the basis of race, color, national origin, sex, age or handicap in admission or access to, or treatment or employment in, its educational programs or activities. Inquiries concerning Title VI, Title IX, and Section 504 may be referred to Wyoming Department of Education, Office of Civil Rights Coordinator, Hathaway Building, 2nd Floor, 2300 Capitol Avenue, Cheyenne, Wyoming 82002-0050 or 307/777-6218 or to the U.S. Department of Education, Office of Civil Rights, Region VIII, 1244 Speer Boulevard, Suite 300, Denver, Colorado 80204.

Table of Contents

Introduction	5
Chapter 1	9
Prevailing Testing Practices: Implications for Learning	
Chapter 2	13
Connecting Learning and Assessment: Research about Learning	
Chapter 3	21
Grades, Tests, and Ordeals: The American Testing Culture	
Chapter 4	37
Roles, Rules, and Rituals in an Culture of Assessment: Blurring the Distinctions	
Chapter 5	49
Making the Transition: The Assessment of Educational Standards and Outcomes	
Resources	59
Wyoming Teachers' Examples of Worthwhile Assessment	60
Selected References for Further Reading:	
Testing and the Testing Culture	79
Assessment and an Assessment Culture	82
The Anthropology and Sociology of Schooling	83
Readings on Teaching and Learning	84
Contrasting a Culture of Testing to a Culture of Assessment..	85
Checklist of Meaningful Assessment	86
Assumptions about Learning	87

Testing, Assessment, and Learning: Invitation to a Discussion

INTRODUCTION

The purpose of this book is to introduce and inform readers about forms of assessment which are consistent with outcomes based education. We hope this book provokes thoughtful discussion about issues important to learners, teachers, and policy makers. We hope it will focus important questions about teaching and learning. How does a book about testing and assessment relate to a Wyoming mandate for outcomes based education? We view these forms as assessment in service to learning.

In our view, assessment in service to learning has four main purposes: a) it includes and informs the student about his or her learning; b) it provides information for the teacher who plans instruction; c) it informs parents about student progress; and d) it provides worthwhile information for policy makers. Our position is unequivocal. The student comes first, the teacher and parents second, and all others third.

This book is divided into five chapters. The first chapter is entitled "Prevailing Testing Practices: The Implications for Learning." The second chapter, "Connecting Learning and Assessment: Research about Learning," is a summary of some of the latest and most pertinent research data about learning.

Chapters 3 and 4 widen the lens to include testing and assessment in a cultural perspective. In these chapters, we show that practices which exist in a culture of testing will probably need to be recast in a culture of assessment.

We get down to the nuts and bolts in Chapter 5: standards for outcomes based education and examples of performance based assessment that complement an outcomes based learning approach. The examples are from elementary and secondary teachers throughout Wyoming. We have identified each because we want to celebrate their ideas. We also want readers who have questions to be able to contact these teachers directly, to continue the discussion started here.

Finally, in a reference section, we have included thirteen examples of worthwhile assessment from Wyoming teachers, and four resource lists for further reading. The book closes with three summary sheets which can be easily duplicated. One summary sheet contrasts the culture of testing to the culture of assessment. A second highlights the characteristics of meaningful assessment. The third summary sheet describes the assumptions about learning that undergird this book.

In part, this book is the product of a teaching experience. As teachers, we think that is fitting. During the spring semester of 1992, we taught four parallel sections of a required course for undergraduates in the teacher education program at the University of Wyoming. The course was Educational Tests and Measurements. Along with colleague Alan Moore, we taught tests and measurements from a classroom assessment perspective. As a result of our efforts, this redesigned course will also have a new name: Classroom Assessment. Throughout the book, we will refer often to the 160 students we taught who want to become elementary and secondary teachers. A formal needs assessment plus ongoing conversations with them about their learning and testing experiences helped shape many of the ideas we present here. Our students helped us remember that many of our teaching beliefs and resulting practices are rooted in the classroom experiences we had as students.

We owe a thank you to many contributors. First, we thank the Wyoming teachers whose assessments are spotlighted in this book. We also thank students in our University of Wyoming courses. Critical friends, Judy Ellsworth, Jane Nelson, and Linda Goldman gave invaluable comments to early drafts. We thank Christine Shearer, whose desktop publishing talents put the document together in what we hope is an interesting and useful format.

Writing a book like this one is a performance assessment of the first order! Throughout the planning, writing, and editing, one question focused our work: How can we invite Wyoming educators to enter a discussion linking testing, assessment and learning?

Audrey M. Kleinsasser
University of Wyoming
Laramie, Wyoming 82071
307/766-6358

Elizabeth A. Horsch
Kelly Walsh High School
Casper, Wyoming 82609
307/577-4640

August, 1992

Chapter I

The Implications of Prevailing Testing Practices

Consider this Scenario... It is the day the teacher has designated for the test. The mood and the atmosphere of the classroom is different from previous days. The students and the teacher are serious and businesslike. Many are tense and anxious. The desks are pushed as far apart as possible. The teacher asks the students to put away all materials, including books, notebooks, examples, models. An empty desk top and a pencil or pen remain. The teacher distributes the test and issues the instructions: "Work carefully and quietly. If you have questions, raise your hand and I will come to you. Keep your eyes on your own work." At the end of the time period which the teacher has designated, students give their tests to the teacher. The teacher identifies the wrong answers, computes the total and records a number or letter which represents the grade in the gradebook. Sometimes, the test will be returned to the student. If the test is a major or standardized test however, the student may not be allowed to see the corrected test.

NOTES & QUOTES

The testing model we have described above remains one of the most enduring, pervasive and compelling features of the enterprise we call school. No other single activity has such a major impact on teachers, students, curriculum and instruction. When preservice teachers from our tests and measurements classes at the University of Wyoming interviewed eighty Wyoming teachers about testing practices at their schools, these in-service teachers expressed major concern about the information

that formal testing produced and the impact of testing programs on curriculum and kids.

Their concerns are well founded. For countless students, tests have become the gatekeeper of their academic journey. Commercial large scale tests, such as the College Board's SAT or the ACT affect college admissions, the awarding of scholarships and financial aid, and the students' perceptions of their own academic abilities. In many states, testing is the sorting, assigning, and classifying system which begins in kindergarten or before. The construction, administration, and taking of tests consume inordinate amounts of teacher and student time. Although there is little agreement about what tests really tell us, the results of formal tests are used to sort and classify students on the basis of academic potential and readiness throughout their school careers. Curriculum decisions, judgments about teacher competence, and comparisons between schools are often made using information from a single, high stakes test.

Wide scale use of test results raises important questions about testing. In this book, we will examine some of these issues about testing, learning, and assessment. What are the impacts of current testing practices on teachers and students? What information does the traditional kind of testing yield? What beliefs do teachers and students have about testing and learning? How compatible are current testing practices with current research about the way students learn? What are the roles, rules, and rituals which limit the opportunities for change? And finally, how are real teachers in real schools changing their assessment practices?

Our purpose in writing this book is to initiate a dialogue about assessment among teachers, students, and the community which can give direction to educators and policy makers as we try to reform education in this state. We invite you to become a part of this discussion and to share your insights and your expertise.

A PRACTICE IN SEARCH OF A DEFINITION

The scenario we have depicted matches Webster's Seventh New Collegiate Dictionary definition of testing. It is "an ordeal or oath required as proof of conformity with a set of beliefs." The key words are "ordeal" and "proof." As we will document in the following section, students are more likely to view tests as an ordeal than as proof of learning. We asked the 160 preservice teachers in our four sections of Educational Tests and Measurements at the University of Wyoming to tell us about their experiences with testing. Their responses testify to the profound and lasting influence of testing. Many expressed powerful feelings about tests. One student considered tests "an inevitable, horrible experience which comes at the end of every chapter or unit." For other students, years later the effects of a particular testing experience were still graphic and traumatic. "In the fourth grade we were finishing a test and when the entire class finished, we could go out for recess early. I was not finished and my friend yelled: 'If these dummies are not finished yet, then they shouldn't get a recess.' I will never forget that."

Other students regarded testing as an ordeal imposed upon them. One said, "This is the way they make you learn the material," while another described a test as, "Panic. Not a fair judgment of one's actual knowledge."

Even those students who did not consider testing hurtful, expressed concern that the tests did not offer reliable proof of what they had learned. Describing his experience with the ACT, one student said, "In taking the ACT, I just chose a blank to fill in if I didn't know the answer....My overall test score was 23 and I did equally well in Math and English, although I hadn't taken Algebra II or Calculus/Trig and having

We must overcome the lazy habit of grading on the curve as a cheap way of setting and upholding standards.

(Wiggins, 1989)

taken and gotten high grades in all the advanced English courses. This demonstrated the incredible inadequacy of the test!"

These comments underscore the impact of testing on more than grade point average. Our students raised two important issues about the nature of testing. Do formalized tests focus on a deficit model of learning? Do tests tell students and teachers more about what a student does not know than what she or he has learned? As one preservice teacher observed, "Is it me, the test, or the teacher?"

The enduring nature of the traditional testing model suggests we have high expectations that testing should yield large amounts of useful information for students, teachers, administrators and policy makers. Calls for educational reform at both state and national levels include a demand for a national or state test to prove that students are learning. Yet it seems clear that all the testing we have done has not given us irrefutable evidence of what our students know or can do as a result of their schooling. To paraphrase an earlier response, *is it the student, the test, or the teaching?*

In the chapter that follows, we examine what the latest research says about learning. How can assessment be connected to human learning?

Chapter II

Connecting Learning and Assessment

The testing scenario we described in Chapter I implies a set of beliefs about the way students learn. Although testing has significant and lasting effects on learning, teachers and students seldom discuss or identify the beliefs which direct the process. Neither teachers nor students are likely to question the purpose of a test. When a student initiates a discussion about a test, the discussion centers on what material will be on the test. Occasionally, the test score or the fairness of the test will be questioned. Testing is so ingrained in the routine of schooling that questioning either the purpose or the methodology is unthinkable. Although teachers and students alike express dissatisfaction with testing and the information it provides, a serious examination of the relationship between testing and learning does not occur.

What beliefs about learning are consistent with the testing scenario we described in Chapter I?

We argue that traditional testing implies a belief that learning is an individualistic activity and must be demonstrated by the student working alone on an individual test. Teachers arrange chairs to emphasize this belief. Students sit formally in straight rows, as far from their peers as possible. The atmosphere is intended to discourage discussion and collaboration. When students share information or compare meaning during a test, this activity is labeled cheating. Even in classrooms where the daily routine includes lively discussion and the exchange of ideas, the test day is silent and solitary.

NOTES & QUOTES

Traditional testing models also suggest that the teacher expects the student to demonstrate or to recognize only those meanings which have been learned previously. It is apparent also that the teacher or the textbook writers identify the meaning. Since talking is taboo and the time frame is teacher prescribed and brief, students cannot contemplate or explore new meaning. The role of the student is to recognize the right response, even if it is cleverly hidden among carefully constructed distractors in multiple choice items, a matching sequence, or in a true/false question.

Students in our teacher preservice tests and measurements classes emphasized this view of testing with these comments. One student identified the test as an instrument to "see how well the student had learned." For most students and teachers the test is not even expected to be a learning experience. It is a demonstration of a previous happening separate from an ongoing, sometimes circular learning process.

Traditional testing also clearly establishes who is in control of learning and who controls how mastery will be exhibited. Teachers give tests; students take tests. The student is the passive question answerer, not the active question generator. As students progress through their years of schooling, they seem to become more willing to assume a passive role. They become so conditioned to giving up responsibility that they resist accepting a more active role. Some educators have observed that by the fifth or sixth grade, students express a preference for activities which demand only rote learning and recall. A colleague shared a disturbing story which illustrates this kind of student detachment. On a particular day in her high school literature class, she mistakenly distributed a test to her students which asked them to answer questions about a piece of literature they had not studied. Although the questions comprised a major part of the test, not a single student questioned the applicability of the test to the materials they had studied. Until she began grading

papers, she did not know that she had distributed the wrong test. Incidents such as this illustrate how important the control issue is and how conditioned students have become to relinquish it without question.

As we examine the traditional testing scenario, we propose that the often repeated admonition to "put away all materials" implies the belief that students can exhibit their knowledge and skills in a tool free environment. Although the importance of "hands-on" learning receives much support among educators, the traditional testing environment is devoid of props and is deliberately tool-free. The instructional setting may have been rich and varied but the testing situation utilizes minimal tools. Although a flutist would not be asked to demonstrate mastery of the flute by naming its parts, students in a chemistry class often take a paper and pencil lab test, and no one questions the practice.

Since a serious discussion of the relationship between beliefs about learning and testing practices does not occur, we cannot be sure that traditional testing actually reflects the commonly held beliefs about student learning. However, we can examine a related question. How do beliefs about learning that are exemplified by traditional testing match current research and thinking about student learning?

We argue, and we think that most educators would agree, that students learn best when they are excited about what they are learning. Students who see themselves as active initiators rather than passive responders learn more deeply and retain knowledge more effectively. When teachers relinquish some control of what is learned and how to learn it and when students are allowed to monitor their own progress, learning is enhanced and accelerated.

We propose that the same standards must be applied to testing. Students need to participate in the selection of the time and the method of evaluation. They must have ownership in both processes. An example of the power of this practice came from

one of our preservice teachers. She described her experience in this way. "In my freshman year in college, I took an anthropology exam which was completely multiple choice and I received a D- grade. I went to my instructor to let him know that I had studied and knew the material and that I was shocked to get a D. He agreed to give me an essay test the next morning and I got a 95%. The remainder of the semester, I took essay exams and finished the course with a high A!"

Making connections. Logic and experience tell us that learning must fit into and expand the framework of what is already known. The previous experiences of the learner and the "real world" applications with which he or she is familiar should be the basis of designing instruction. A student who cannot identify a connection or link a new idea with an existing concept has little chance of making sense of the idea. Mary Budd Rowe, former President of the National Science Teachers Association who has spent many years studying how students learn science, uses an elevator analogy to illustrate this point. Once the idea is on the elevator, it must first get off on the right floor. When the elevator door opens, the idea pours down the hall. If the door to a room where the idea belongs opens, then the idea finds a place to stay. If not, it just continues down the hall, down the stairs and is lost forever.

A student must construct knowledge using what he or she already knows. Traditional tests which measure bits and pieces have severe limitations in this respect. Test questions generally are brief. The teacher provides little or no context information. Discussion is not allowed. Thus, the student may not identify the framework of the idea. Faced with this dilemma, most students who have not "gotten off on the right floor" may resort to blind choice and guessing.

The context of learning. The idea of context is closely related to making connections. Although there are differing views on the subject, some researchers

propose that the skills and ways of thinking which make a discipline unique are acquired in the context of the content and methodology of that discipline. It appears unlikely that a set of generalizable skills exists which can be transferred easily from one discipline to another. Classroom teachers have struggled with the manifestations of this concept for a long time. They observe that students who exhibit proficiency in writing in a writing class have difficulty writing coherent lab reports in science, or that the mathematical skills which students display in a math class do not transfer when they are asked to figure their test percentage in a social studies class. The deficit may lie in the student's failure to recognize the context rather than in the lack of his or her skill. Similarly, some observers postulate that the dramatic increase which students exhibit on the College Board's SAT or the American College Testing Program's ACT scores after completing the practice course probably is the result of improved ability to recognize the context, rather than the result of any real gains in learning.

We accept the premise that learning is difficult in a tool-free atmosphere. It is reasonable to assume then that the tools which a student uses become an integral part of his or her knowledge base. The books, calculators, manipulatives, and a host of other instructional devices become so entwined with learning that students may find it difficult to demonstrate learning without them. This belief requires a very different testing environment from the standard "paper and pencil and nothing else" model which prevails in most testing situations. Although teachers exhort students to consult the dictionary and reward them for citing multiple sources on a research paper, these same students may be expected to prove what they know without any tools on the test.

Teaming. Increasingly in the world of work, teams work together to solve problems and make decisions. Educators also are beginning to acknowledge that

expanded learning is more likely to occur when students have the opportunity to discuss, probe and explore meaning in groups. When a student tests ideas against the judgement of others, he or she is more likely to develop higher order thinking skills. Rote memory and recall may flourish in an solitary setting but exploration, synthesis and analysis are enhanced by collaboration.

As teachers, we acknowledge the difficulty of writing tests which probe high level thinking skills. Perhaps the problem lies more in the methodology of testing than in the testing instrument. It may be that a solo performance severely limits the opportunities for exhibiting thinking. Recalling a cooperative final, one of our preservice teachers said, "I took a final in an education course with a group of six people. We openly discussed the questions. I learned more than I ever had just taking a test alone."

Another research finding which has major implications for teaching, testing, and learning deals with the concept of learning as a hierarchy of lower and higher level thinking skills presumed to occur in some sort of predetermined progression. Our observations convince us that children can organize and synthesize in remarkable ways. As very young children develop language skills, they demonstrate this capability. Yet schools spend vast amounts of time teaching basic skills and having students practice them before students are challenged to apply these skills in meaningful ways. Such an educational process may be unnecessary and even detrimental to the development of high order thinking. Tim Rush who teaches at the University of Wyoming cites work done with learning disabled children. He observes that it is not necessary for a child to become absolutely competent in the skill of sounding the letter "p", and to be required to exhibit this facility repeatedly, before he or she can learn to say "Peter Piper picked a peck of pickled peppers."

Brain research. Research on brain development may offer support for equal and coinciding emphasis on higher order skills and basic skills. The areas of the brain which support memory and the sites for complex thinking are different. Learning theorists postulate that these areas develop independently and that increased emphasis on one kind of learning at the expense of the other may adversely affect the rate of development of the less used area. Although the research is tentative, the implications are not trivial.

Summary. Some of the beliefs about learning can be summarized as follows:
Learning...

- depends on the active involvement of the learner in interpreting as well as recording information and in monitoring his or her own knowledge construction strategies
- depends in part on the learner's ability to place an idea or skill within the framework of what is already known
- occurs in context. In real life learning, tools, props, and other aids are often a part of the context in which one learns.
- is not automatically transferred from one context to another. There are probably severe limitations on generalized learning abilities.
- occurs most successfully in a social context. Knowledge that is posed, exposed and tested with others is more likely to have real and lasting meaning for the learner.
- proceeds from the general to the particular. Mastery of detail should not be expected before the big picture is explored.

These emerging beliefs about learning deserve serious consideration. Assessment and instruction should be integrated. Testing can no longer be considered as an isolated event which always occurs as an interruption or

culmination of the instructional process. When testing does occur at a designated time, it will reflect methods of instruction and it may even occur when the student decides he or she is ready to demonstrate proficiency. Assessment strategies will call for the use of tools, self monitoring, multiple answers and evolving ideas. Students will exercise some control over the form of assessment. They will find and design many ways to exhibit their achievements. Teachers will not limit the evidence for learning to a single bit of information gathered at a particular time by a particular instrument. Parents and the community will accept multiple and less structured indicators as "proof" that children are learning.

Teachers, students and the community will share responsibility for assessment and instruction. The current atmosphere of structured, prescribed tests of knowledge will give way to messy and varied indicators of achievement which mirror the way real people learn about real problems.

Chapter III

Grades, Tests, and Ordeals: The American Testing Culture

In this chapter, we look at testing through the lens of culture. We argue that many American testing practices thrive within a cultural framework which includes roles, rules, and rituals for teachers and students. If educators are to implement alternatives to traditional paper and pencil tests, it is important to understand the culture that supports this tradition. We will explore the idea that a testing culture is unlikely to be hospitable to worthwhile assessments such as portfolios and other forms of performance based assessment.

NOTES & QUOTES

EXAMINING SCHOOLING FROM SOCIOLOGICAL AND ANTHROPOLOGICAL PERSPECTIVES.

In the last 15-20 years, it has become commonplace for some educators to examine schooling using a sociological or anthropological perspective. A brief description of the impact sociology and anthropology have had on the way educators view schooling follows.

Contributions of sociology. A sociological perspective maintains that schooling is larger than the individual student or teacher. To better understand a group, sociologists often examine the relationship of socio-economic status, age, race, gender, religion, and family structure. Sociologists have also studied the role of teachers, the school as a work place, and teaching as a profession. One example of

sociology's explanatory power relates to understanding change or the lack of change. Sociological studies offer insights about the reasons why the institution of school seems to resist change. Teachers and administrators at all educational levels are bombarded by calls for restructuring, renewal, and reform. In the view of the cynics among us, calls for various forms of restructuring, not to mention plans for cooperative learning, site-based management, and outcomes based learning simply represent this year's new thing which replaces last year's new thing.

A second example of the way sociology informs educational research and practice relates to a community's expectations for its students and schools. A community communicates expectations for academic achievement in various ways. Some of the expectations are subtle and covert, others are overt. For example, it used to be commonplace for nine weeks and semester honor rolls to be printed in school and community newspapers. The 1990s version of support for academic achievement takes a completely different format in some communities. The message that formerly appeared in newspapers now flashes from the rear bumpers of cars and trucks: I'M PROUD TO BE A PARENT OF AN HONOR STUDENT.

A third example of sociology's power to help us understand school involves the values honored by a community. One example is the way a community may or may not support sports. In a community that does not value soccer, it will not matter how skilled the soccer coach is or how much money the school board allocates to the program. Soccer is not likely to thrive in that culture.

Contributions of anthropology. Research about schooling has also benefited from the perspective anthropology offers. The anthropological perspective is holistic. It includes human beliefs, practices, and organizational patterns. Until recently, most anthropologists were expected to travel to remote locations to study groups of people with different languages from their own,

distinctive family and community roles, and unique rules of behavior which produced cultural rituals. However, some anthropologists came to see that there was much to be learned at home, much to be learned about uniquely American cultures and sub-cultures. This awareness has led researchers to examine American schooling from an anthropological lens.

The result of an anthropological study of education has yielded a finding that would not surprise many American students or teachers. Namely, educational anthropologists view formal American school practices as a rich and complex culture with its own roles, rules, and rituals. Studies using anthropological methods of interview and observation over long periods of time have led some researchers to believe that rural schools differ from urban schools, and elementary schools differ from secondary schools. More important, anthropological studies suggest that there can be great variations among schools classified as either rural or urban. Factors such as ethnicity, religious values, and economic status of the community create important differences between communities that may be of similar population size. In fact, particular anthropological studies suggest that some of the common adjectives used to describe schools and schooling, for example, a white, middle class school, may not be meaningful or even useful, especially to teachers. Schooling is just more complex than the labels used to describe it. Whenever teachers try to address issues about school climate and culture, they are acknowledging that effective teaching and learning is complex, and not easily generalizable from community to community, or even from school building to school building within the community.

Educational anthropology has spawned a relatively new research field called critical theory. One of the characteristics of critical theory is that teachers become researchers in their own classrooms. Teachers tell their own research "stories."

Ecological psychology was able to examine "big picture" differences between big schools and small school.

(Barker & Gump, 1964)

Critical theory examines the importance of race, gender, and ethnicity in a formal learning setting. Studies of this type have informed pre-school education and literacy. For example, communication rules in some cultures demand that children avert their eyes when an adult speaks to them. When a teacher has this information and acts on it, a child who might have been described as uncooperative, even belligerent, is now more accurately understood to be respectful of a teacher.

The implication of this anthropological finding for an instructional approach such as cooperative or collaborative learning is stunning. How could a child who is taught that adults represent a superior authority easily think of his or her teacher as a collaborator, as an equal partner in learning?

As described above, sociological and anthropological research studies have informed educational practice and theory. These studies challenge oversimplified solutions to educational problems. Many oversimplified solutions are being proposed at the national or international level with respect to large group, standardized testing. Oversimplified solutions do not address the roles that students and teachers have in a school culture, the school rules that teachers, students, and others follow, or the schooling rituals most of us take for granted.

In the remainder of this chapter and the one that follows, we contrast the culture of traditional forms of testing to a culture of assessment. By understanding that practices like testing occur in a cultural setting, educators can unravel the meaning of rules, roles, and rituals which have defined American testing practices throughout this century.

American schools' culture of testing. American schooling perpetuates and exists within a culture of testing. "When's the test?" is usually one of the first questions a student asks the teacher. It is a good question. For most students and teachers, the school year is highlighted by testing events that have become

For a specific example about eighth grade students, see Atwood's (1987): In the Middle: Writing, Reading, and Learning with Adolescents.

traditional. Some of the testing events can be characterized as "high stakes" testing. Consider the following examples of high stakes tests:

- placement tests for special education;
- Chapter I screening tests;
- individual intelligence testing for placement in gifted education programs;
- quarterly, mid-term, and semester tests;
- fall and spring achievement testing;
- state-mandated achievement testing;
- college entrance examinations like the College Board's SAT.

Usually, the implementation of high stakes testing is not in the control of classroom teachers. High stakes testing does not make a difference in most students' day-to-day classroom life because they fail to see the connection among the tests, curriculum, and instruction. Students do not realize that a curriculum or a set of text books may have been adopted in their school because the results of traditional standardized testing indicate deficits. Policy makers hope that a different text book and related curriculum materials will address the deficits. The results of high stakes testing also lead to significant decisions controlled by policy makers at the state and national levels: college admission, the creation or elimination of special programs.

In addition to high stakes tests, students are bombarded with low stakes tests. These are tests that occur more frequently. Often, they are teacher-made and driven by the curriculum. Examples of low stakes tests include the following:

- unit test;
- chapter test;
- test over the book;
- test over the material;
- end-of-the week test;
- pop quiz;
- lab test.

There is ample evidence that most American students are being tested regularly with as much precision as state-of-the art standardized test procedures demand and human administration allows. Admittedly, teacher-made tests are less precise. This lack of precision is what testing experts refer to as weak reliability and validity. Yet, despite traditional testing, some researchers maintain that American students are the most under-examined students in the world. By under-examined, they mean that students are not asked probing questions to explain and substantiate their answers. Nor are students required to self-appraise their performance. Despite all of the high and low stakes testing in schools, traditional testing practices do not tell students, teachers, or parents very much about the way kids learn or how they think about what they have learned.

The predictable cycle of calls for mastery learning, or now, performance based assessment and portfolio assessment underscores the point that current testing practices do not seem to reveal what students have learned or need to learn. Some sources suggest that students spend six weeks a year in testing events. Moreover, current testing practices bear little relationship to the way people learn. See Chapter 2, "Connecting Learning and Assessment" for a more complete discussion of current learning theories and why we believe the boundaries between assessment and instruction are blurred.

Clearly, high and low stakes testing have an effect on students and teachers. For most students, *any* test, even a five point quiz, is a high stakes test since there is something to lose or gain: confidence, self esteem, the respect of the teacher. As discussed throughout this book, specifically, the comments from undergraduates who want to become teachers, the effects of traditional testing and grading are more than psychological. Traditional testing affects friendships and group membership, for example, honor societies and extra-curricular activities. In the section that follows,

we describe some of the roles, rules, and rituals that teachers and students live by and observe in a culture of testing.

ROLES, RULES, AND RITUALS IN A CULTURE OF TESTING

The teacher's role. In a culture of testing, the role of the teacher is clear. The teacher writes tests or uses commercially prepared tests. The teacher monitors the test setting to discourage cheating and writes down the results in a grade book. Some teachers enter numbers into a micro-computer, using a software program to calculate percentages. It is the teacher who explains the results to students and parents. In a culture of testing, the teacher's role is one of total control and responsibility for classroom testing. Yet, most teachers feel uncomfortable in the role of judge or evaluator. Recent research points out that if they had a tests and measurements course in their teacher preparation program (and not all teachers have this training), for a majority, it was a course they neither liked nor enjoyed. Thus, most teachers are in the uncomfortable position of knowing that they are supposed to be an evaluator at the same time that they lack the confidence to be a competent evaluator.

Part of a teacher's role is to keep tests from becoming contaminated. Contaminated tests are those for which the questions or answers have somehow gotten into the hands of students or others. Contaminated tests are no longer reliable or valid. In a culture of testing, one of the rules is that only the test giver knows the questions and answers. Such a test is known as a secure test. From philosophical and theoretical standpoints, "teaching to the test" or making sure that students know what is going to be on the test, is a major point of debate among testing experts. Yet, there is something unsettling about the contradiction of requiring students to attend

Gullickson's 1986 study reveals that many teachers do not feel competent as evaluators of student work.

school for the purpose of mastering skills we as a culture have decided are important to learn (cultural skills and knowledge) and then devising ways to keep tests about those skills secret. Keeping tests uncontaminated or secure is one way that a culture of schooling makes testing mysterious to students and even to some teachers. The mystery is part of what makes traditional testing powerful.

The chief role of the teacher in a testing culture is that of a judge, evaluator, and scorekeeper. In a culture of testing, a culture demanding accountability at local, state, national, and international levels, teachers and students are well aware of test scores that symbolize failures, deficits, or mistakes. Teachers traditionally report points gained or percentages earned from a test. When a teacher is a scorekeeper, not a competent evaluator of student work, what he or she is reporting is a deficit. Thus, instead of celebrating the success of learning, a culture of testing emphasizes what was *not* learned. Some readers may disagree with our interpretation of a testing culture that calculates and reports student deficits rather than celebrates successes. However, we contend that celebrations like the ones observed in a testing culture come for too few students, too infrequently.

The testing culture produces a kind of academic economy. Some researchers have described grades as the coin of the realm. Instead of nickels, dimes, and silver dollars, the currency includes A's, B's, C's, D's, and F's. The academic economy hinges on the theory that there must be fewer A's than B's or C's. After all, if everyone received A's and B's, newspapers would not publish honor rolls. One of the myths of grading (especially at the postsecondary level) is that a course with a majority of A's must be an easy course, evidence of grade inflation. When we think about grades as an economy, it is another way of understanding that in a culture of testing, only a small percentage of students will have the right to celebrate success as measured by grades. Typically, we label these students as gifted or academically superior. Other

winners in an academic economy are a school's valedictorian, salutatorian, or national merit scholarship winners. Our point here is *not* to lessen the accomplishments of academically successful students. We simply point out that in the economics of A-F grade distributions, there can only be a few winners, perhaps as few as four or six percent receiving A's. The A's represent education's wealthy.

In summary, the role of the teacher in a testing culture is the person in control. He or she is in control of the test, the test site, management of the test results, and the economy of grades. The teacher is an evaluator, and may be unconfident in an evaluator's role. The student rarely challenges. But a teacher's authority is localized. The higher the stakes of the test, the less control the teacher has, although traditional standardized testing affects decision-making about curriculum and textbooks, which in turn affects classroom instruction. In this country, the control of a high stakes tests like the College Board's SAT rests with an independent body. One example is the Educational Testing Service in Princeton, New Jersey. High stakes testing like that represented by the SAT, the American College Test (ACT), or the Graduate Record Exam (GRE) is a topic undergoing intense debate at national and international levels. While the reliability of these tests goes unchallenged (they are highly reliable), their validity remains controversial: do results of a the GRE predict anything other than the results of a similar kind of test? The purpose of this book is not to address the high stakes testing debate or the helplessness and lack of ownership most teachers feel in relation to them. For readers interested in the topic, we have included sources for further reading.

The student's role. A culture of testing clearly indicates what the role of the student is. For students like the ones we surveyed at the University of Wyoming, testing is in their words, "inevitable," sometimes an "ordeal." Words like inevitable

and ordeal suggest that the student has little or no control over the testing situation, and in fact, often feels like a powerless victim.

Students rarely schedule tests or other kinds of exit performances for themselves. One exception to this is what is known as "challenging the first chair" in music. This is how it works. If a trumpet player thinks she is capable of sitting first chair in the section (a position musicians usually acknowledge as the best player and leader of the section), she may ask the director for the opportunity to challenge the current first chair player. The challenge is based on an authentic performance of the musician's skill *when the musician knows the time is right*.

However, in most school cultures, students are examined when teachers "need a grade for the grade book." Typically, students are examined at a natural stopping point in the schedule. Stopping points are often calendar driven, not curriculum and outcome driven, i. e., tests at the end of the week, just before a major holiday, and at the end of the quarter or semester. One of the unique characteristics of most traditional testing situations is that the student does not determine *when* he or she is ready to be tested or on *what*. A second characteristic of traditional testing is that it tends to "conclude" ideas. For example, when the test over the Revolutionary War is completed, so is most discussion about the Revolutionary War.

Students in a testing culture can be mystified by a grading process that teachers control, maintain, and manage. One of the ironies of a testing industry that produces micro-computer software which will generate percentages, stanines, interquartile ranges, class averages, and student averages, all at the touch of a key on a micro-computer keyboard, is that most students do not know why a teacher gives them the grade they get. Research tells us that students who perform the poorest in school are also the least able to accurately appraise why they are receiving the

grades they are receiving. Where do grades come from and what do they mean? Too many students simply do not know.

Rules of testing. Just as there are distinct roles for teachers and students in a testing situation, there are distinct rules of testing. One testing rule in American schools is the "do it alone" rule. A student completes most tests by him or herself. Sometimes, "doing it alone" means an absence of calculators, microcomputers, dictionaries, periodic table charts, or any of the tools characteristic of everyday work in fields like economics, business, and chemistry.

A second rule of most testing events in schools is the dominance of traditional, paper-pencil tests. Paper-pencil tests are inexpensive to create, efficient to administer, and except for essays, they are scored objectively. Typically they take the form of true or false, multiple choice, completion, and short answer. Sometimes paper-pencil tests are used even when they do not make much sense, for example, a paper-pencil test to determine the grade in a physical education course or a word processing course. Using paper-pencil tests when a performance or a demonstration would be more meaningful is one more piece of evidence that the rules of testing are part of a powerful school culture, a culture students and teachers seldom question or challenge. We suggest that the predominance of paper-pencil tests is the result of a lack of confidence teachers have as competent and qualified evaluators.

A third rule of most testing situations is that the teacher evaluates the test. This makes the teacher the sole audience for the test. One of the outcomes of the teacher being the audience for the test is that students give up ownership for their work. Who does the test belong to? The questions and the answers belong to the teacher.

We asked 160 preprofessional teachers enrolled in an educational tests and measurements course at the University of Wyoming why university students fail to pick up their end-of-the-semester exams from the instructor. Some of their responses follow:

All these students are concerned with is their final grade. They don't care how they got it, just so they got it. They feel it will not do any good to come pick up their test, it will not change their grade.

Students no longer are worried about the final exam grade after they receive their final grade. The grade on the exam has little relevance after they receive their report card.

Students often only worry about one semester at a time. Once that semester is over, it's over, and the student starts fresh. When receiving a course grade that a student is satisfied with, the student "blows off" the paper or exams he has taken.

Because they found out their overall grade from their report cards and felt it was right. They didn't see what it would matter because their grades were done and there was nothing they could do about it.

... there is nothing they can do to change it (good or bad).

All of the responses above share a common theme about ownership and responsibility. The test is not the student's. It is not worth picking up. Picking it up will not make a difference in the student's grade. The only thing that counts is the grade on the report card. The responses also suggest that students fail to make connections between their courses and learning.

Rituals of testing. The rituals of testing in American schools are easily identified. Usually, tests come at the end of a unit of study. Sometimes (as in a pre-test or diagnostic testing for special needs students) they occur on the very first day

of school. Often, the test is different from the instruction that preceded it. For example, the lecture and group discussion that precedes a history exam may be followed by a paper-pencil test that the student completes alone. Most of the students we surveyed at the University of Wyoming experienced and viewed instruction as being different from testing. This important point is one we will return to in our discussion of the development of an assessment culture and the blurring of instruction and assessment.

A second ritual of testing is the synthesis and pulling-together that students preparing for tests engage in. Students may review all of their notes, write summaries of important pieces of information, complete sample problems or essays. As noted previously, in a testing culture, synthesizing is an end-of-the-unit exercise, not an ongoing learning activity. The test signals closure and conclusions. One characteristic of synthesizing is that students stop to think about what they have learned or what they need to learn. In this way, testing literally punctuates the school year by slowing down or interrupting learning. As teachers, most of us have probably experienced the sinking sensation that students "memorize for the test." These same students may fail to make sense of the knowledge so that it can be meaningfully applied in non-testing settings. This is one of the outcomes of a testing ritual that is summative rather than formative and ongoing. However, as teachers, we must also acknowledge that traditional forms of testing encourage memorizing for the test and de-emphasize connection-making.

A third ritual of testing involves mental and physical preparation. How many times have you, in your teacher role, solemnly and conscientiously directed students to "get a good night's sleep, eat a good breakfast, come prepared mentally and physically." If such instructions are given only on a test day, they send a clear message to students about the importance of tests in contrast to ordinary school

activities. Such messages may support a student's suspicion that a test is an ordeal requiring extraordinary stamina!

A culture of testing has roles, rules and rituals for students and teachers. The roles, rules and rituals of a testing culture are summarized in the figure on the opposite page. Cultures do not change easily. Sociologists and anthropologists have found that groups of people continue doing what they do because roles, rules, and rituals are purposeful. They represent long-held and accepted traditions. In a testing culture, one purpose of testing is to sort and weed students. The culture of testing is a powerful one in American schools, supported by tradition and a testing industry that can produce reliable and cheap paper-pencil tests. However, moving from a testing culture to an assessment culture is not impossible. In the following chapter, we describe what an assessment culture looks like and where it already exists.

Culture of Testing

ROLES

- teacher is the sole authority, sole audience for a test
- teachers keep tests uncontaminated, so they can be used over and over again

RULES

- student usually has to go it alone
- testing is a zero sum game (competition)
- letter grades are produced, this is a deficit report, "economy of thrift"
- grades create a classroom economy predetermined number of As, Bs, Cs, Ds and Fs
- dominance of paper/pencil tests which carry more value than anecdotal records
- test giver (sometimes the teacher) is sole evaluator
- knowledge is disaggregated into separate pieces of knowledge (e.g., math, reading, social studies, science) so it can be easily tested via p/p testing

RITUALS

- learning for the test and only the test, e.g., "Will there be a test over this?"
- test signals the end of "school learning," strategic forgetting, dumping
- testing is sometimes different from the instruction that precedes it
- testing is mysterious - where do grades come from?
- testing is an ordeal, a test of fire
- testing is a red flag word for most of us
- the process is tradition and familiar to most, thus it is a process that is seldom questioned.

Chapter IV

Roles, Rules, and Rituals in a Culture of Assessment: Blurring the Distinctions

In the previous chapter, we suggested that roles, rules, and rituals can be identified in a testing culture. What we will describe next is an assessment culture. Assessment does not carry the emotional baggage that the word testing carries. Airasian (1991) defines assessment as:

the process of collecting, synthesizing, and interpreting information to aid classroom decision making; includes information gathered about pupils, instruction, and classroom climate.

It may be important to note that some people use the terms assessment, testing, evaluation, and measurement interchangeably. For educational researchers and psychometricians, each term has a separate and technical meaning. In this chapter and throughout the book, we are using Airasian's definition of assessment. A culture of assessment looks and feels different from a testing culture. Assessment blurs the roles, rules, and rituals that seem to be distinct, perhaps unchangeable in a culture of testing. We begin with the roles of teachers and students.

Teacher and student roles. A key characteristic of assessment is bringing students into a conversation about their own learning. Another way of thinking about this idea is to ask students to take an active role in their own learning and its assessment. Teachers move toward collaboration with students when teachers and students share assessment decisions. Many teachers demonstrate this ideal when

NOTES & QUOTES

they instruct students to "Read four books of your choice from this list of ten." On an essay test, a student may be instructed to "choose three of the four essays." A teacher may provide three or four options to a project a student must complete.

While the examples above suggest that a student has some choice about what to complete, the responsibility and authority of creating the questions and assessing the products remain in the teacher's hands. In an assessment culture responsibility and authority are shared by the student and the teacher--for questions and answers. The student shares ownership of the assessment. At a fundamental level, students and teachers share ownership of what learning is in a content area, for example. Together, students and teachers grapple with hard questions like the following: a) What does it mean to know something? b) What is the relationship among instruction, curriculum, and assessment? c) How do outcomes for particular content areas relate to one another?

In a testing culture, the teacher's assessment role is clear: the teacher is the final and only appraiser of an acceptable answer or product. Appraising means evaluating the worth, significance, or status of their work. In an assessment culture, students are self-appraisers and co-appraisers with teachers and other competent evaluators. With training and a healthy critical perspective, students take an active role in the assessment process. Thus, the traditionally distinct roles of teacher and student are blurred in an assessment culture.

Portfolio assessment offers one example of student appraisal and ownership. First, some of the materials for the portfolio are selected by the students. Selected materials are based on criteria. Selection is purposeful. Second, in their portfolios, often in an introductory statement, students explain why they have included their choices. These statements are valuable insights to parents and teachers about a student's intentions. They tell what the student cares about, why the student thinks it

is done well. Perhaps more importantly, by selecting some of the materials that go into the portfolio and justifying their inclusion, students own their work: students have some control over portfolio contents. This is much different from the secure, uncontaminated test that belongs to the teacher or another test giver in a testing culture. What we are describing is a collaboration based on solid criteria.

By appraising their own work, students learn to have opinions and make informed, thoughtful judgments. In a culture of testing, only the test giver appraises, has opinions, and makes judgments. It may be more accurate to say that the only judgment that counts in the end is the one made by the test giver. However, the effects of the testing may have long lasting impact on the test taker.

One of the most difficult tasks of moving from a testing culture to an assessment culture is convincing students that self appraisals are a necessary part of an assessment culture. Students are being asked to change what is a well-established student role. It is likely to be particularly difficult for junior high and senior high students who are newly introduced to the ideas of portfolios or performance assessments. Students and teachers will need training and peer support for an assessment culture to thrive. For students and teachers to take the risks inherent in the move from a testing culture to an assessment culture, safe environments for risking must be established. It will require parent information sessions. It will be helpful for teachers who are uncomfortable with the idea of students sharing assessment responsibilities to spend some time talking about collaboration and role definitions. Teachers should elicit student questions and responses. It may be wise to start moving into a culture of assessment slowly, perhaps in one assignment, rather than revamping the entire course. Another idea is to begin with less rigorous assessment. For example, many teachers are doing this already when they ask students to complete drafts of written work over the period of a semester or even a

year. Collecting the drafts, final copy, and evaluation criteria for portfolio inclusion is one way to begin meaningful assessment collaborations between students and teachers.

The rules of an assessment culture. We identified three rules in a testing culture: going it alone, the predominance of paper-pencil testing, and the teacher as a single audience. In the section below, we look at the way these rules are bent, literally reshaped, in an assessment culture.

As we have emphasized in earlier chapters, going it alone (working alone on a test), seems to be a characteristic of a school setting, not the real world. In fact, working cooperatively and collaboratively is sometimes considered cheating in a school setting and frowned upon as being less demanding. Business and industry have long challenged educators to prepare students for the world of work. A current challenge to schools is to prepare students for the greatly changing work place of the twenty-first century. Working cooperatively is one skill needed for the future. Other skills include a capacity for self reflection and for sustained engagement in a task. Higher order thinking skills are being emphasized as never before.

In a culture of assessment, students are collaborators and work cooperatively. By cooperatively, we refer specifically to cooperative learning as conceptualized by Johnson and Johnson, and Slavin. Cooperative learning requires interdependence and individual accountability. Sometimes, cooperative learning may involve peer evaluation.

In an assessment culture, a second rule of testing is bent. It is the use of paper-pencil testing. In an assessment culture, in addition to or sometimes in place of paper-pencil tests, students complete what some experts are calling worthwhile assessments. Worthwhile assessments are also referred to as alternative assessments, authentic tests, and real tests. The term performance based assessment is yet another

way of describing a worthwhile assessment. In a performance based assessment, students demonstrate what they know and can do in a real-life or meaningful situation.

Performance based assessments are nothing new or even revolutionary. One example most of us are familiar with is a driver's license procedure. To be licensed, the driver must complete two tests, both worthwhile, both different. First the applicant must complete a paper-pencil exam about rules of the road at a particular proficiency level. Even with a perfect score on the exam, the license will not be granted without a behind-the-wheel driving examination conducted by an examiner who asks the applicant to perform skills important to driving safely and with a minimum of skill. Since every driver is unique, so is every behind-the-wheel exam. Based on the examinee's level of experience, the examiner makes an informed judgment about licensing the driver. It is important to note that there is little disagreement that both exams are worthwhile. It is also important to note that the driving exam occurs in a context. If an applicant lives in New York state, the exam occurs there. There is not much argument that people accustomed to rural driving know that city driving has a different set of demands, even though the rules of the road are standardized.

Lastly, in a culture of assessment, the third rule which is bent is that of audience. The teacher is not the sole audience of an assessment. Earlier, we referred to the results of a survey we asked students in our Tests and Measurements courses at the University of Wyoming to complete. Students told us that they do not pick up tests because they do not own them. Picking them up will not make a difference in their grade anyway.

In a culture of assessment, students have an authentic audience for their work, whether it be writing, scientific experiments, or history projects. Many

teachers already demonstrate that they value authentic audiences for student work. Most schools organize and support plays, debate tournaments, art exhibitions, band concerts, vocal concerts, open houses to display industrial arts projects, science fairs, history fairs, all variety of sports competitions and exhibitions. Sometimes the purpose of the event is a competition, sometimes it is an exhibition. Each is a celebration of student achievement, progress, and accomplishment. Not one is organized to exhibit a deficit.

Students develop expertise when they have an opportunity to share the products of their work with an appropriate audience. Students share their ideas for reactions, revisions, or assistance. Audience can include other students, teachers with expertise in that particular content area, parents, and oneself. In a testing culture, there is only one member of the audience the test giver, the teacher. When students share their learning products with other students, the audience broadens. The wider the audience, the more useful the information a student is likely to receive about his or her performance. The teacher cannot and should not be the sole audience for all assessments. However, the teacher should be a member of the audience.

There is an important difference between the displays described above and traditional paper-pencil tests that are not worthwhile assessments. It is unusual for a science project to be left behind in the school auditorium, unusual for a sculpture to be abandoned after an art exhibition. Real scientists and real sculptors would not do that. However, as the education community embraces the idea of worthwhile assessments, it must guard against the trivialization of fairs, exhibitions, and competitions. When a science project is left behind, it is a clear statement that the student did not view the science fair as any kind of celebration of his or her work.

Rituals of assessment. We identified three rituals prevalent in a culture of testing: a test signifies the completion of learning; a test is an ordeal requiring extra sleep and nutrition; and a test is mysterious. The rituals of assessment blur traditional testing rituals.

In a culture of assessment the borders separating instruction and testing are blurred. What some writers call kidwatching and others call informal assessment illustrates how instruction and assessment coexist. Kidwatching has its origin in the whole language movement and studies of emerging literacy (learning to read and write). Most elementary teachers are well-acquainted with the term. Kidwatching is the process of interviewing and observing students to make informed judgments about student learning and instructional planning. The result of kidwatching is not a formal letter grade or external report.

Early in the 1980s, a preschool special education specialist related the following real event to one of the writers of this book. The anecdote shows the power of interviewing as an assessment tool. The specialist was testing three year-olds as part of preschool special education and kindergarten readiness screening. For one test, the preschoolers were supposed to tell her the names of the colors. When the specialist showed a little boy the color green, he readily identified the color as "John Deere." Given the right/wrong nature of the test, the child gave the wrong answer and was penalized for his deficit knowledge: he did not produce the right answer which was "green." However, what would have happened if the specialist had questioned him about his reasons for giving that answer? What if she would have asked a question or two about where he lived, what his mother and father did? What if she would have asked him what John Deere was? Probing questions like these might have given the child an opportunity to demonstrate that his information was contextually accurate. In fact, he could have demonstrated that he knew far more

than the test question asked or measured. In the real world of farming, the little boy knew that green meant John Deere. Just as important, in the course of interviewing, the specialist could have helped the boy make the connection between the concept "green" and the concept "John Deere." Interviewing students is one way to teach and assess at the same time. The assessment event is a learning event. The event blurs the arbitrary distinction between instruction and testing.

A second ritual that is blurred in an assessment culture is changing testing practices from an ordeal into a celebration for both the student and the teacher. The 1980s and early 1990s were not a time of celebration for the givers or the takers of large scale, standardized tests in the United States. Newspaper headlines regularly jolted readers with reports of declining math and science scores and the poor performance of American students. American students fared poorly when compared to Western European and Asian students. The testing frenzy and its soul mate, accountability, have so affected the American school culture it is unlikely that any report of student progress would cause Americans to react with a smile and a "Well done!"

We return to the purpose of this book to emphasize this important point. In our view, in a culture of assessment, the boundaries between learning and assessment are blurred. Meaningful learning coupled with worthwhile assessment includes the student in an ongoing conversation about his or her understanding and growth. Secondary purposes of assessment include providing information for the teacher who plans and modifies instruction and informing parents about student progress.

Third and last, in a culture of assessment, the rituals of testing and grading are demystified. One of the reasons testing and grading is so powerful is that they are both mysterious. Where did they come from? What do they mean? Why is it that

grades are inadequate predictors of future success on the job? It is a mystery to students, especially to those who do not perform well in school.

In an assessment culture, students know why they are being examined and what their performance means. Students know the assessment criteria. Part of the mystifying nature of tests is the language experts use: averages, percentiles, stanines, bell-shaped curves, grading on the curve. All of these words have precise, technical meanings. Testing experts feel comfortable and understand the language of measurement. But sometimes test givers and other experts hide behind the language of testing when they talk with students about their performance on school-related tasks. Administering the test is only half of the job; interpreting the test is the other half.

In an assessment culture, students and teachers share the language of assessment. The following example illustrates our point. How does a reader assess a story? Can a third grader assess a story? Can a third grader identify the characteristics of an effective story? Is the story interesting, boring, funny, sad, true-to-life? This is the language of assessment in the real world of books and literacy. There is nothing mysterious about it. It is not owned by the teacher. It is not too complicated for a third grader to talk about with another third grader, a teacher, a parent, or even another writer. The roles, rules, and rituals of an assessment culture are summarized in the figure on page 47.

Conclusion. In this chapter we contrasted an assessment culture to a testing culture. We conclude by repeating one of the points we made early in the chapter. Cultural practices are resilient. Cultural practices do not change easily or quickly. Teacher and student roles in a testing culture are clearly defined and the product of long held traditions that define the American school. Because we (educators and noneducators) have been produced by that culture, most of us are comfortable with

the clear boundaries provided by the roles, rules, and rituals of a testing culture. They tell us what to do and how to do it. Redesigning the roles of teacher and student in an assessment culture will not be easy. Worthwhile assessments are time consuming to produce, validate, and administer. Worthwhile assessments represent a different kind of evaluation paradigm. It will be tempting to try to fit new assessment needs into familiar and traditional evaluation paradigms. The two cultures are contrasted at the end of the book, on page 85.

The expectations in an assessment culture parallel the changes American culture has experienced with respect to family life, community life, and work place. The cultural roles, rules, and rituals for the twenty first century demand a different kind of teacher and a different kind of student. We believe that it will demand a culture of assessment rather than a culture of testing. In the next chapter, we focus on transitions, from testing to assessment.

Culture of Assessment

ROLES

- student as worker, as the expert, as a competent self evaluator
- student as generator of knowledge, not just a receptor of knowledge
- student and teacher share roles of authority and/or learning
- teaching is building connections (more than facilitating)
- teaching/learning is a conversation, not telling

RULES

- collaboration, teaming, cooperation (adult/student, student/student)
- an assessment usually has multiple audiences (self, peers, teachers, others)
- worthwhile assessment takes many forms: paper/pencil, performance, demonstration, exhibition
- we can have more than one winner, learning isn't a zero sum game
- rigorous subjective judgment that is data based

RITUALS

- worthwhile assessments are a celebration not an ordeal, it is/can be public, not just private in the classroom or school
- ongoing synthesis, ongoing questions and conversations about "Is this making sense?"
- rough drafts and journaling are legitimate pieces of evidence that learning occurred
- assessments are more than paper/pencil evidence: play, competition, demonstration, contest, open house, science fair

Chapter V

Making the Transition: The Assessment of Educational Standards and Outcomes

Consider this Scenario...It is a normal day in the classroom. The prevailing mood is serious and intent. Students work in small groups to finish specific tasks. They confer with each other and between groups to coordinate their work. The room is filled with books, maps, computer equipment, and other materials. On the wall, the students have posted an outline of the tasks to be completed and the assignments for each small group. A timeline shows when the parts of the project must be completed to meet the deadline for presentation of the final report to an outside group. The teacher circulates quietly among the groups, responding to questions and monitoring group dynamics of the groups. Near the end of the period, the teacher and the class discuss the progress that has been made and the projected work for the remainder of the project. As a group, they define and assign responsibility for the necessary tasks. The teacher distributes a checklist for each student which lists some of the identified learner outcomes including the knowledge and skills which are embedded in the work of the project. She asks for feedback on the checksheet and reminds the students that they will be responsible for setting goals and monitoring their own progress toward mastery of the outcomes on the checksheet. Although the students are interested and attentive, they clearly regard this process as a part of the routine of the classroom. The conversation quickly switches back to the project and the preparations for making the results of their work available to the community. As

NOTES & QUOTES

Since 1970, scores (of U.S. students) on basic skills tests have been steadily increasing while scores on assessment of higher-order thinking have been steadily declining in virtually all subject area.

(Darling-Hammond, 1991)

"Outcomes are nontraditional, future-oriented abilities students will need to be productive citizens of the 21st century.

Our district has taken on the mission of graduating students who are 1) self-directed learners, 2) collaborative workers, 3) complex thinkers, 4) quality producers, and 5) community contributors.

[Outcomes] become our curriculum, the focus of our instruction, and eventually our graduation requirements.

But to know when students have achieved the Five Outcomes and to be able to document that achievement required a radical change in assessment.

(Redding, 1992)

the work session ends, the students return their papers and materials to their working files and the teacher completes her anecdotal records on her clipboard.

Do these activities constitute legitimate testing? Certainly there is little evidence to link this scenario with the testing scenario described in Chapter One. In a testing culture, we have been engrossed in measuring bits and pieces of student knowledge. As we make the transition from a testing to an assessment culture, we first must be able to decide what we want students to know or be able to do. At the state and national level, we are struggling to reach a consensus and the focus is shifting from behavioral objectives to meaningful outcomes as the standards of student achievement. Outcomes, as they are beginning to be defined at state level, reflect a very different view of what students should know and be able to do. These outcomes indicate an awareness of the knowledge, skills and attitudes which students will need to function effectively in an increasingly complex and technological society. As the community takes a more active role in defining goals for the school, the definition of what constitutes appropriate education is changing also. Merely succeeding in school is not enough. Schools and communities are expecting students to become more than information storers. They are articulating the view that the "learnER" is more likely to succeed than the "learnED", that lifelong learning implies the ability to direct one's own learning in purposeful and productive ways.

Although some schools and communities have identified unique student outcomes, there are many commonalities among the characteristics they cite as essential. The ability to solve complex problems, to access and critically analyze information, and to adapt to change are identified time and again as critical skills for all students. Social and interactive skills, self awareness, and the ability to question one's own position often make the list. Perhaps the most interesting and significant idea which appears among the list of expectations is the emerging view of students as

immediate, active community contributors and an untapped resource for improving their own lives and the lives of others in the community.

If schools are to be accountable for assessing students' achievement of these complex outcomes, what kinds of assessment procedures will be valuable? How will the results be reported? How will students, teachers and parents make the transition from reliance on a single grade to multiple indicators (observations, interviews, and performances of real life tasks) as evidence that learning has occurred?

To find answers to these questions, we interviewed Wyoming teachers. Their names, addresses and strategies are detailed in the People Resources section (page 59 of this book). As these skilled and thoughtful teachers shared their experiences with us, we identified some common and some unique concerns among teachers who are designing and piloting alternatives to traditional testing in their classrooms. These teachers had a clear vision of the knowledge and skills they considered essential for their students. They were also confident of their assessment practices. However, many of them found it difficult to translate their observations into a record of progress --a report-- which could be accepted and understood by the public.

For some teachers, the attempt to massage the information gained from these ongoing and informal kinds of assessment activities into more standardized measures distorted both the activity and the evaluation. Dana Van Burgh, a middle school earth science teacher, described a summative performance assessment in his classroom which he used to finish a unit on maps, rock types, structures, and elevations. Each team of students was given a black and white Landsat image of the county. Teams labeled as many features as possible, including streams, towns, roads, and geologic features. To perform the task, the teams had to work from a variety of maps with different scales and different ways of presenting information. His delight in the abilities of his students to process information and to use bits and pieces to create a

Students who have to
perform or exhibit their
own knowledge and
skills get learning in the
bones: active learners
become lifetime
learners.

(Rexford Brown, 1989)

composite picture was obvious. When we queried him about how he translated the success of the students into numbers in the grade book, he remarked that simply giving the students credit for the number of landforms they could identify seemed a sterile and insufficient measure of their achievement. He worried aloud about three summative tests grades which he had recorded for each student and the disparity between the students' test scores and their ability to utilize data and to solve problems he had observed as they were engaged in the activities of the classroom.

Other teachers expressed similar concern. Vicki Foster, whose sixth grade class engaged in a year-long study of the effects of an oil spill on a local stream, supplied us with a list of the kinds of performances she used to determine the progress of her students as they squared off against a real life problem. Her students drafted a personal conservation ethic, wrote and presented speeches to the community based on their findings, designed wall murals to display their data, and engaged in debates on controversial issues. She used no objective testing and in her words, "The grades I assign to students are hard to justify to someone who does not observe first hand what the kids can do."

We found teachers pioneering unique ways of reporting student progress also. In a first grade classroom in Yoder, Wyoming, Judy Lissman has replaced report cards with portfolios as the method she uses to report to parents. She keeps a record of each child's activities and his or her performance on specific skills on a video tape. This video tape, along with other portfolio items, is shared with parents during the quarterly parent conference sessions. Parents view the tape, offer comments, and respond to questions about their child's progress. Judy commented that an unexpected dividend from the taping activity has been the awareness the children develop about their own progress and their interest in charting their improvement.

We have a strong tendency to proclaim the educational poverty of test scores and then turn around and use them, when we can, as indices of our own success, thus legitimating the validity of the public concerns about the quality of education. If test scores in their conventional form do not reveal what really matters in school, we should not use them to judge our "success."

(Eisner, 1991)

Riverton Elementary School engages the student, the parent, and the teacher in monitoring student achievement in an ongoing process. Once or twice a week the student is asked to write a letter home, explaining a particular learning activity. Wayne Dennis, Principal/Teacher of the school states, "These letters help us to know if the students truly understand the process we are teaching, whether in math, science, social studies, or reading."

Teachers are also embracing forms of informal assessment. Diana Wiig, who teaches first grade in Rock Springs, reports that assessment at the primary level in her school is informal. She describes her own assessment protocol in this way, "Children come to school very excited about learning. Testing seems to dampen their enthusiasm. Testing students at an early age tends to tell us only what they DON'T know, not what they know. My assessment begins on day one. Since young children are always eager to show you what they know, that's where I begin. We sit down and talk, one-on-one. With forty students, that can take a lot of time, but my partner and I deem it well worth the time spent. We also observe the children at independent learning activities and in group situations. We keep a notebook handy to write down our observations. If a child concerns us, we may start an anecdotal file and add more details."

The Natrona County School District's Lighthouse Park School has replaced grades with an individualized, ongoing assessment project. There the teachers have compiled a skills checklist based on such process skills as "the ability to consult a variety of sources," "understands that the reading must make sense," and "shows an appreciation for the environment." The marking key for the skills checklist defines the categories of Beginning, Developing, Independent, Commendable, or Not Assessed this Quarter. Records are kept on the completion of assignments and the quality of the work done. Each child has a portfolio and video tapes of presentations or special

This mode of
assessment (portfolios)
conjures a different
perception of a
classroom—a portfolio
cultural center where
knowledge is
constructed
communally, individual
strengths are
developed and learning
becomes an adventure
where success is
displayed.

(Collins, 1990)

projects. A personal computer disk is the repository for evidence of the child's language skills. At the end of the year, the teachers confer with parents to determine what materials will remain in the portfolio to become a part of the child's permanent record. Doris Waddell, the teacher of the third/fourth grade in the Lighthouse school, reports that the teachers find this system of assessment both exciting and challenging.

For some teachers, the use of projects as a form of assessment creates interesting insights about student learning. Greg Hammonds, who teaches in a middle school, uses a product testing project as an assessment of his students' mastery of process skills. After a preliminary classroom exercise in which the students compare various brands of paper towels, the students choose a different commercial product and compare different brands. They must design and conduct their own experiment. Hammonds' criteria for evaluating the projects is based on data collecting skills, the validity of the project design, data analysis, and how well the conclusions are supported by the data. One of his concerns was the difficulty students had in applying the methodology and skills learned in one context to another problem. For him, this adds another complexity to the problem of designing performance assessments which are novel but also accurately reflect what students have learned.

In other cases, the use of alternative forms of assessment (not paper-pencil) to assess student progress toward identified outcomes has had significant impact on the curriculum. In Worland, as a result of their experiences with outcomes and alternative assessment, the middle school teachers recommended the adoption of a curriculum which de-emphasizes any formal testing. In Torrington, Bob Pesicka, an elementary school principal, noted that what the teachers learned about their students from more thoughtful assessment caused major science curriculum change.

We found the process of creating performance assessment symmetric with good teaching activity to be time consuming, requiring considerable scientific and technological know-how.

(Shavelson & Baker, 1992)

Teachers in that elementary school devoted considerable time to student interviews in an ongoing effort to find out precisely what students knew and were learning as a result of their instruction. Based on the information and insights this kind of assessment produced, teachers discarded or revised parts of the K-5 curriculum. They put criterion referenced testing on hold. According to Pesicka:

What we have experienced is to see the top down model (outcomes, curriculum, lessons, assessment) becoming part of a development cycle so that what we learn about the learner drives the process. Our intent is to respond to the learner results as observed by the teachers to determine when and how our outcomes and curriculum need adjusting.

Although the methods these teachers used to assess student learning were different, some common elements were obvious. Through these conversations and based on the writings of other educators, we have identified some of the characteristics of meaningful assessment. We offer this checklist as a starting point for designing and evaluating alternative assessment.

The main feature of dynamic assessment techniques is that they seek to evaluate students' potential for change by either observing their response to instruction or determining the processes responsible for change.

(Campion, 1991)

Characteristics of meaningful assessment...

- is based on a clearly articulated consensus of what the learner must know or be able to do;
- demands that the learner construct meaning as well as respond to questions;
- probes the process used to answer the question;
- assigns tasks that are both meaningful to the learner and accurate measures of the instructional program;
- allows for more than one acceptable answer;
- probes understanding of the big picture;
- allows for alternative ways of demonstrating achievement;
- includes the use of appropriate tools;
- allows for individual and group performance;
- includes self and peer evaluation and student-teacher collaboration in designing the assessment process;
- is an ongoing process which measures growth;
- is a celebration of success rather than proof of failure;

Our search for examples of alternative assessment in Wyoming gave us some definitive answers to the question of what good assessment looks like. However, as teachers and schools make the transition from testing to assessment, how the results should be reported remains an unanswered question. Informal assessment is a powerful tool for verifying that students have learned but the data come to us in a disaggregated form which may not be acceptable to policy makers and the public. Meaningful assessment does not yield a single number which can be conveniently bubbled in on a data sheet. Yet the demand for meaningless quantification may not allow the time and freedom necessary for new techniques and methods of assessment to mature into acceptable measures of student achievement.

As we continue the dialogue about testing, learning, and assessment and as the unanswered questions continue to plague us, perhaps we will do well to remember that good questions are better than poor answers. We have been mesmerized by the quantity of data test givers can extract from a single testing episode, but what these data tell us about student learning remain meager.

The good news is that performance assessments do not duplicate information about student achievement provided by traditional tests (average correlation is about 0.45). They tap different aspects of achievement, but we are not sure what aspects of achievement multiple choice test or performance assessment do or do not tap.

(Shavelson & Baxter, 1992)

Resources

EXAMPLES OF WORTHWHILE ASSESSMENT

How to Access Information in This Section

The resources on the following pages highlight strategies Wyoming K-12 teachers use to meet instructional outcomes through worthwhile assessments. Each example includes the teacher's name, school address and phone, the grade level in which they use the activity, and a brief description of the activity.

NOTES

VIDEO ASSESSMENT

CONTACT

Judy Lissman
Southeast Elementary
Yoder, WY 82244
532-7176 (work)

GRADE LEVEL

First Grade

CONTENT

Description of the activity: During the year, video tapes of each child engaged in a variety of activities are made. Tapes of field trips, birthdays and other special events are recorded also. The video record is correlated with a portfolio method of reporting to the parents. The portfolio also contains paper and pencil samples of the student's work.

Formal and Informal Use of Assessment Information: The portfolio is shared with the parents quarterly. The video is sent home with the parents for viewing and a questionnaire is included to solicit parent response. The teacher views the tapes as a method of becoming more aware of student behavior and for noting specific areas which are targeted for improvement. Students also view the tapes and assess their own performance.

BOLTON CREEK STUDY ASSESSMENT**CONTACT**

Vicki Foster
Jackson Hole Middle School
P.O. Box 568
Jackson, Wyoming 83001
733-4235 (work)

GRADE LEVEL

Sixth Grade

CONTENT

Description of Activity: The sixth grade class conducted a year long study of a stream which had been impacted by an oil spill. The study was done in cooperation with the Bureau of Land Management and the Wyoming Game and Fish Department.

Assessment activities included:

1. Students wrote a personal conservation ethic at the end of the unit. The writings were evaluated to see if the students had personalized the information.
2. Students wrote and presented speeches about the project to various community groups and to other classes in school.
3. Students reported results of the study either pictorially or in writing.
4. Role plays were used.
5. Students debated the controversial issues after researching.

Formal and Informal Use of Assessment Information: No objective tests were given. (Teacher's comment: It is hard to justify the grade which was given to someone who does not observe what the students can do. I hate report cards.)

PRODUCT TESTING PROJECT

CONTACT

Greg Hammons
Worland Middle School
1200 Culbertson Ave
Worland, WY 82401
347-3233 (work)

GRADE LEVEL

Seventh Grade

CONTENT

Description of Activity: Students engage in an instructional activity in which they compare the properties of various brands of paper towels. The activity is a review of the science process skills. After this teacher directed activity, each student is required to choose a product and design and conduct an experiment to compare different brands of the chosen product.

Formal and Informal Assessment Information: The project is evaluated by the teacher based on (1) data collection, (2) critical thinking and data analysis, (3) project design, and (4) whether the data supports the conclusion. The grade on the project becomes a part of the student's final course grade. The science teacher works with the language arts teacher during the project. Students prepare a "commercial" as a means of reporting their conclusions and product test results. The commercials are video taped and evaluated by the language arts teachers and the science teacher for both content and presentation.

LIGHTHOUSE SCHOOL ASSESSMENT PROTOCOL

CONTACT

Doris Waddell
Park Lighthouse School
140 W. 9th Street
Casper, WY 82601
577-4593 (work)

GRADE LEVEL

Third/Fourth Grade

CONTENT

Description of Activity: A portfolio is kept for each child. Examples are included from all subjects and pieces are selected to show progress in a particular skill area. Teachers use a skills checklist which they have compiled. Such skills as "the ability to use a variety of resources," "understands that reading must make sense", and "shows an appreciation and respect for the environment" are included. A key which is used to mark progress on skill acquisition uses the categories of Beginning, Developing, Independent, Commendable and Not Assessed. Teachers keep a gradebook and record completion of assignments, quality of work done, and when a skill is mastered. A notebook containing anecdotal records on each child is maintained also. Special projects or presentations are recorded on video tape.

Formal and Informal Assessment Information: No grades are given. At the end of the year, the parents, child and teacher make decisions about what should remain in a child's portfolio for the next year.

FIRST GRADE MULTIDISCIPLINARY ASSESSMENT

CONTACT

Diana Wiig
Washington School
Rock Springs, WY
362-3411 (work)

GRADE LEVEL

First Grade

CONTENT

Description of Activities: Assessment is informal and ongoing. A variety of tools are used. Depending upon the information/performance being sought, individual observations are used to chart progress in targeted skill areas. Content is integrated and projects are used for assessment. For example, the first and sixth grades collaborate on growing plants. The students record data, write in a plant book, and compare rates of growth. They write a story about how it felt to be a seed and what the seed was thinking as it waited to sprout.

Alternative Assessment Examples

CONTACT

Wayne Dennis, Principal/Teacher
Riverton Intermediate School
121 N. 5th St.
Riverton, WY 82501
856-7852 (work)

GRADE LEVEL

Fourth and Fifth Grade

CONTENT

Description of Activities: Reading assessment: Student interviews and student journals are used. In the interviews, teachers check for comprehension skills (how well the students uses context clues, previous information), re-reading techniques and personal knowledge. Key questions are used to assess these skills. They include:

- a) what does the passage talk about?
- b) what key words helped you know that?
- c) where did you get that information?
- d) did you re-read the text to get the information?

The journals are used to assess such skills as comparing, contrasting and recognizing plot, fact and opinion.

Presentations: Students in Social Studies relive the past by becoming a person born in 1700 or 1800, learning their trade and building a town. Presentations are made to parents and other students. Knowledge of time period, oral communication skills, research skills and writing skills are assessed by this kind of activity.

Projects: Students build a colony, produce video and cassette tapes, write books or stories using a particular period in history.

Letters: Students write letters home to parents explaining what they have learned in a certain subject.

Performance Outcomes: Outcomes are just being written. Outcomes and rubrics for effective mathematics have been completed. These are self, peer and teacher evaluated.

MONOGRAPHS

CONTACT

Dana Van Burgh
Dean Morgan Junior High School
1440 S. Elm
Casper, WY 82601
577-4440 (work)

GRADE LEVEL

Ninth Grade Enriched Classes

CONTENT

Description of Activity: Students select a topic of interest to them and the community. They do the research (which may include interviews and locating of existing photographs), write the text and produce camera ready layout for publication. Topics can be selected from any discipline. Examples of monographs which have been completed include Casper's First School, The Old City Hall, Sticks and Bones, and The Sand Dunes.

Formal and Informal Assessment: The monographs provide insight into the student's research techniques, persistence and ingenuity. The writing, after revision, is an example of their best work. The selection and use of photographs and the layout is an indicator of artistic effort.

RENAISSANCE FAIR

CONTACTS

Barbara Matteson, D'Arcy Patey
Natrona County High School
930 S. Elm Street
Casper, WY 82601
577-0355 or 577-0330 (work)

GRADE LEVEL:

English Literature Honors, English Literature Advanced

CONTENT

Description of Activity: English Literature Honors and English Literature Advanced are part of the 11th grade core curriculum at Natrona County High School. Each spring, these classes hold a Renaissance Fair as the culminating activity for the their study of the English Renaissance. The Fair provides an opportunity for students to be immersed in the culture of the period. For example, students select an area to research in depth and present to visitors as part of a vignette. Vignettes cover various aspects of English life. For example, one group focuses on a Shakespearean play and cuts it for a 20 minute presentation. The players would also research acting so that they could explain to visitors about the kind of lives actors led. Visitors include grade school students.

Other groups learn about different aspects of English life. Often this involves cooperation across the curriculum, because students who are madrigal singers need the help of the choir director. Those who are presenting English country life work with the science teachers to learn how to make essentials like soap and candles. Often outside resource people will teach students about a specific skill or craft.

Who uses the assessment information? The assessment of student performance is judged by the English literature teachers and is both formal (students receive a grade) and informal (information will be used to plan for next year's fair). The assessment is shared between teacher and student or groups of students in a debriefing. The debriefing focuses on the value of what the students learned about the Renaissance and how effective the students were at teaching what they learned to audience members. Comments from fair visitors help confirm the teacher's evaluation.

Formal and Informal Use of Assessment Information: The information is used formally (grades) and informally (fine tuning the fair for next year). In the oral debriefing, teachers talk with students to determine if knowledge gaps exist about the period. If there are gaps, they may be taught traditionally. The formal grade is based on the level of involvement and the demonstration of understanding, actual participation.

SELF EVALUATED GYMNASTICS

CONTACT

Charles Zurbuchen
Wyoming Center for Teaching and Learning - Laramie
Box 3374 University Station
Laramie, WY 82071-3374
766-2155 (work)

GRADE LEVEL AND CONTENT

Eighth and Ninth Grade Gymnastics

CONTENT

Description of Activity: Two groups of eighth and ninth grade physical education students participated in a gymnastics unit which was composed of seven stations or apparatus. Each class was divided into seven groups of three to four students. Students had one week to practice at each station and were evaluated the following week.

Olympic style judging (0-10) was used and criteria were established for each apparatus by the instructor. The first period class was evaluated solely by the teacher, while in the second period class, the students evaluated each other. The first period required three class sessions for the teacher to evaluate each student, therefore the students' activities at each station had to be reduced. The second period evaluations were done in two periods, providing ample time for students to retest activities at each station.

Who uses the assessment information? The assessment information is used for informal assessment and useful during parent teacher conferences. Comments by students who favored peer evaluation included the following: it was interesting to learn how to score people in gymnastics and to be scored by my

friends; grading each other was fun as long as we all were honest about it; it was hard to grade each other and I think we grade tougher than the teacher would have; we had more fun taking the test than if the teacher was there; when we did the scoring, it was more like a class activity than a test.

Formal and Informal Use of Assessment Information:. Assessment and instruction are blurred in this activity.

INTEGRATION OF COMPUTERS

CONTACT

Wayne Schatz
Central Middle School
25 South Custer
Sheridan, WY 82801
672-9014 (work)

GRADE LEVEL

Sixth and Seventh Grade Integrated Computer with Core Area

CONTENT

Description of Activity: The computer teacher and the language arts teacher (this year's focus) team teach computer lessons. The sixth grade teachers encourage students to build a portfolio of written work to be shared with the seventh grade language teachers. Each student has been writing on the computer with the program Bank Street Writer III. Each child has his or her own disk on which to save all stories. At the end of this school year, each student chose one of their stories and used the program, "Ghost Writer" produced by the Minnesota Educational Computer Consortium (MECC) to analyze the story. The results were printed and kept in the student's portfolio.

"Ghost Writer" does the following: a) readability test; b) sentence length analyzer, c) clarity check, and d) vocabulary analyzer. The program gave comments on the student's writing. Students looked at the summary and wrote any comments on it that they wanted to share. The printed summary was added to the student's portfolio and given to next year's language arts teacher.

Who uses the assessment information? The goal is to assess writing every spring. The information is given to the next year's teacher so that she or he

can begin skill assessment immediately, thus avoiding a "dry assessment." The assessment is also a conference tool between student and teacher.

Formal and Informal Use of Assessment Information. For the present, the assessment is informal, going into a temporary writing file that follows the student. Stories selected by the student and the teacher are included as well.

LITERACY ASSESSMENT IN THIRD GRADE

CONTACT

Patsy Tomb
Coffeen Elementary
1053 S. Sheridan
Sheridan, WY 82801
674-9333 (work)

GRADE LEVEL

Third Grade Literacy

CONTENT

Description of Activity: Teachers developed a form to be used during small group discussion. Students keep a record of what students were doing.

First, students discuss the book. When students and teacher decide that everything's been covered, they jot down important points that would be included if they were to write a summary. Each of the points is assigned a point value according to a consensus of the group.

The teacher determines points for the reading section by asking student if they have read the assignment. If the teacher has doubts after the discussion, s/he can ask the group what it thinks about each person's score. Students do not try to get points when they aren't earned, once they understand that the whole group will be responsible for assessing the work. Frequently, students assess more harshly than the teacher, never the other way around. Students and teacher agree on a rule that the teacher has the last say in the final points awarded. Students have approached the teacher after class to say that they didn't deserve all of the behavior points because they were messing around during group time.

Who uses the assessment information? Together, teacher and students use the information.

Formal and Informal Use of Assessment Information. See the assessment rubric below.

CHECK LIST

Reads Assignment (5 points)

Journal

Activity

Discussion

Behavior (10 Points)

Comments

READING AND WRITING IN FIRST GRADE

CONTACT

Bette Sample
Gertrude Burns Elementary School
116 Casper Avenue
Newcastle, WY 82701
746-4671 (work)

GRADE LEVEL

First Grade Reading and Writing

CONTENT

Description of Activity: Each day, by noon, first graders in my classroom have a masterpiece that they've written and read to me, other students, or another adult (the principal or reading consultant). This past year, our school put a moratorium on standardized tests in the primary grades (K-2). My goals, in regard to assessment, are to determine the progress made by each first grade student, every day. I have an informal assessment team. Most of the members do not know they are participating. The team is made up of colleagues, administrators, parents, the students, and me. Students read and write each day. They share their work with their classmates, my colleagues, our administrators, and their parents.

Who uses the assessment information? At the end of the school year, I create a portfolio of their writing (four-five pieces) for inclusion in the cumulative folder. Next year, I plan to collect something at the end of each month. Their work ranges from two to three sentences to four pages. Their writing speaks for itself, I do not include a teacher narrative about it.

ASSESSING THE INTEGRATION OF READING, WRITING, SOCIAL STUDIES AND SCIENCE

CONTACT

Joyce Tyrrell
Meadowlark School
550 S. Burritt
Buffalo, WY 82834
684-9518 (work)

GRADE LEVEL

Second Grade, Interdisciplinary

CONTENT

Description of the Activity: The goal of this assessment is to assess district skills objectives as students use them in higher level performances that correlate to a district outcome for interdependence. The specific skills sheets were used by the teacher to check for district skills objectives. The rubrics were used by the student as a guide for the end product and were used by the student and the teacher to assess the final product.

Students made a poster including the objectives on the rubric. They wrote a three paragraph paper explaining their poster (based on the three sections of the rubric) for the purpose of teaching a kindergarten student about birds and insects. They also had the kindergarten student make an insect out of clay. The second rubric is the delivery rubric--how well they thought they taught their kindergartner. The teacher then marked off district objectives based on the rubric, the student's paper, and their delivery to the kindergarten student.

Who uses the assessment information? The audience for the assessment includes the students, the teacher, the principal, and the parents.

Formal and Informal Use of Assessment Information. The assessments of this activity are informal for the present. They do not go to the third grade teacher. While the assessment does address a district outcome, accountability for meeting the outcome is the district's next step.

Sources for Additional Reading

TESTING AND TESTING CULTURE

- Airaisian, P. W. (1988). Symbolic validation: The case of state-mandated, high-stakes testing. Educational Evaluation and Policy Analysis, 10(4), 301-313.
- Airasian, P. W. (1991). Classroom Assessment. New York: McGraw-Hill.
- Anderson, B., & Phipho, C. (1984). State mandated testing and the fate of local control. Phi Delta Kappan, 66(3), 209-212.
- Atkin, M., Kennedy, D., & Patrick, C. (1989). Inside Schools: A Collaborative View. New York: Falmer Press.
- Brown, R. (1989). Testing and thoughtfulness. Educational Leadership, 46(7), 31-33.
- Chittendon, E. (1991). Authentic assessment, evaluation and documentation of student performance. In V. Perrone, (Ed.) Expanding Student Assessment. Alexandria, VA: ASCD.
- Collins, A. (1990). Portfolios for assessing student learning. In A. Champagne, B. Lovitts, and B. Calinger, (Eds.) Assessment in the Service of Instruction. (pp. 157-166) Washington, DC: AAAS.
- Cuban, L. (1984). How Teachers Taught - Constance and Change in the American Classroom, 1890-1980. New York: Longman, Inc.
- Darling-Hammond, L. (1991). The implications of testing policy for quality and equality. Phi Delta Kappan, 13(3), 220-225.
- Eisner, E. (1991). Should America have a national curriculum? Educational Leadership, 49(2), 76-81.

- Ember, C.R., & Ember, M. (1981). Cultural Anthropology. Englewood Cliffs, N. J.: Prentice-Hall, Inc.
- Goodman, Y. (1985). Kidwatching: Observing children in classrooms. In A. Jaggar & M. T. Smith-Burker (Eds.) Observing the Language Learner. (pp. 9-18). Newark, Delaware: International Reading Association.
- Gullickson, A. R. (1986). Teacher education and teacher-perceived needs in educational measurement and evaluation. Journal of Educational Measurement, 23 (45), 347-354.
- Hudson, L. (1990). National initiatives for assessing science education. In A. Champagne, B. Lovitts, and B. Calinger (Eds.) Assessment in the Service of Instruction. (pp.61-79). Washington,DC: AAAS.
- Jaeger, R. (1991). Legislative perspectives on statewide testing-goals, hopes and desires. Phi Delta Kappan, 13(3), 239-242
- Johnson, D. W., & Johnson, R. T. (1974). Instructional goal structure: cooperative, competitive, or individualistic. Review of Educational Research, 44, 213-240.
- Johnson, D. W., & Johnson, R. T. (1987). Learning Together and Alone. Englewood Cliffs, N.J.: Prentice-Hall.
- Koretz, D. (1992). A Rand Note. Evaluating and Validating Indicators of Mathematics and Science Education. Santa Monica, CA: RAND.
- Madaus, G.(1991) The effects of important tests on students: implications for a national exam system. Phi Delta Kappan, 73(3), 226-231.
- National Commission On Testing and Public Policy. (1990). From Gatekeeper to Gateway: Transforming Testing in America. Chesnut Hill, MA: Boston College.
- Redding, N. (1992). Assessing the big outcomes. Educational Leadership, 49 (8), 49-53.
- Resnick, D. P., & Resnick, L. B. (1985). Standards, curriculum, and performance: A historical and comparative perspective. Educational Researcher, 14, 5-21.

- Shavelson, R. & Baxter, G. (1992). What we've learned about assessing hands-on science. Educational Leadership, 49(8), 20-25.
- Shepard, L. (1989). Why we need better assessments. Educational Leadership, 46(7), 4-9.
- Slavin, R. E. (1983). Cooperative Learning. New York: Longman.
- Slavin, R. E. (1987). Grouping for instruction in the elementary school. Educational Psychologist, 22, 109-127.
- Smith, M.L.(1991). Meanings of test preparations. American Educational Research Journal, 28(3), 521-542.
- Wiggins, G. (1989). A true test. Toward more authentic and equitable assessment. Phi Delta Kappan, 70(9), 703-713.

ASSESSMENT AND AN ASSESSMENT CULTURE

- Airasian, P. W. (1991). Classroom Assessment. New York: McGraw-Hill.
- Arter, J. A., & Spandel, V. (Spring, 1992). NCME instructional module: Using portfolios of student work in instruction and assessment. Educational Measurement: Issues and Practice, 11(1), 36-44.
- Frederiksen, J. R., & Collins, A. (1989). A systems approach to educational testing. Educational Researcher, 18(9), 27-32.
- Haney, W., & Madaus, G. (May, 1989). Search for alternatives to standardized tests: Whys, whats, and whithers. Phi Delta Kappan, 70(9), 683-687.
- Goodman, Y. (1985). Kidwatching: Observing children in classrooms. In A. Jaggar & M. T. Smith-Burker (Eds.) Observing the Language Learner. (pp. 9-18). Newark, Delaware: International Reading Association.
- Mehrens, W. A. (Spring, 1992). Using performance assessments for accountability purposes. Educational Measurement: Issues and Practice, 11(1), 3-9, 20.
- Stiggins, R. J. (1991). Assessment literacy. Phi Delta Kappan, 72(7), 534-539.
- Stiggins, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. Journal of Educational Measurement, 22 (4), 271-286.
- Wiggins, G. (May, 1989). A true-test: Toward more authentic and equitable assessment. Phi Delta Kappan, 70(9), 703-713.
- Wolf, D., Bixby, J., Glenn, J. III, & Gardner, H. (1991). To use their minds well: investigating new forms of student assessment. In G. Grant, (Ed.). Review of Research in Education, 17, 31-74. Washington D. C.: American Educational Research Association.

THE ANTHROPOLOGY AND SOCIOLOGY OF SCHOOLING

- Barker, R. G., & Gump, P. V., (1964). Big School, Small School. Stanford, CA: Stanford University Press.
- Jackson, P. W. (1968). Life in Classrooms. New York: Holt, Rinehart & Winston.
- Kirschenbaum, H., Simon, S., & Napier, R. (1971). Wad-Ia-Get? The Grading Game in American Education. New York: Hart.
- Lightfoot, S. (1983). The Good High School. New York: Basic Books.
- Lortie, D. (1975). Schoolteacher: A Sociological Study. Chicago: University of Chicago Press.
- Peshkin, A. (1986). God's Choice: The Total World of a Fundamentalist Christian School. Chicago: University of Chicago Press.
- Waller, W. (1932). The Sociology of Teaching. New York: Wiley.
- Wolcott, H. F. (1975). The Man in the Principal's Office: An Ethnography. Prospect Heights, IL: Waveland Press, Inc.

READINGS ON TEACHING AND LEARNING

- Anderson, J. (1983). The Architecture of Cognition. Cambridge, MA: Harvard University Press.
- Anderson, J. (1882). Acquisition of cognitive skills. Psychological Review, 89, 369-406.
- Atwood, N. (1987). In the Middle: Writing Reading and Learning with Adolescents. Portsmouth, N.H.: Boynton/Cook Heineman.
- Caine, R. N., & Caine, G. (1991). Making Connections: Teaching and the Human Brain. Alexandria, VA: ASCD.
- Jacques, E. (1985). Development of intellectual capability. In F. Link(Ed.). Essays on the Intellect. (pp. 107-142). Alexandria, VA: ASCD.
- MacLean, P. (1987). A mind of three minds: educating the triune brain. In 77th Yearbook of the National Society for the Study of Education. Chicago, IL: University of Chicago Press.
- Resnick, L. (1987). Education and Learning to Think. Washington, DC: National Academy Press.
- Resnick, L. (Ed.) (1989). Knowing, Learning and Instruction: Essays in Honor of Robert Glaser. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rowe, M. (1973). Teaching Science as Continuous Inquiry. New York: McGraw-Hill.
- Vygotsky, L. (1978). Mind in Society. Cambridge: Harvard University Press.

Culture of Testing	Culture of Assessment
ROLES • teacher is the sole authority, sole audience for a test • teachers keep tests uncontaminated, so they can be used over and over again	ROLES • student as worker, as the expert, as a competent self evaluator • student as generator of knowledge, not just a receptor of knowledge
RULES • student usually has to go it alone • testing is a zero sum game (competition) • letter grades are produced, this is a deficit report, "economy of thrift" • grades create a classroom economy predetermined number of As, Bs, Cs, Ds and Fs • dominance of paper/pencil tests which carry more value than anecdotal records • test giver (sometimes the teacher) is sole evaluator • knowledge is disaggregated into separate pieces of knowledge (e.g., math, reading, social studies, science) so it can be easily tested via p/p testing	• student and teacher share roles of authority and/or learning • teaching is building connections (more than facilitating) • teaching/learning is a conversation, not telling RULES • collaboration, teaming, cooperation (adult/student, student/student) • an assessment usually has multiple audiences (self, peers, teachers, others) • worthwhile assessment takes many forms: paper/pencil, performance, demonstration, exhibition
RITUALS • learning for the test and only the test, e.g., "Will there be a test over this?" • test signals the end of "school learning," strategic forgetting, dumping • testing is sometimes different from the instruction that precedes it • testing is mysterious - where do grades come from? • testing is an ordeal, a test of fire • testing is a red flag word for most of us • the process is tradition and familiar to most, thus it is a process that is seldom questioned.	• we can have more than one winner, learning isn't a zero sum game • rigorous subjective judgment that is data based RITUALS • worthwhile assessments are a celebration not an ordeal, it is/can be public, not just private in the classroom or school • ongoing synthesis, ongoing questions and conversations about "Is this making sense?" • rough drafts and journaling are legitimate pieces of evidence that learning occurred • assessments are more than paper/pencil evidence: play, competition, demonstration, contest, open house, science fair

MEANINGFUL ASSESSMENT

- is based on a clearly articulated consensus of what the learner must know or be able to do;
- demands that the learner construct meaning as well as respond to questions;
- probes the process used to answer the question;
- assigns tasks that are both meaningful to the learner and accurate measures of the instructional program;
- allows for more than one acceptable answer;
- probes understanding of the big picture;
- allows for alternative ways of demonstrating achievement;
- includes the use of appropriate tools;
- allows for individual and group performance;
- includes self and peer evaluation and student-teacher collaboration in designing the assessment process;
- is an ongoing process which measures growth;
- is a celebration of success rather than proof of failure.

LEARNING...

- depends on the active involvement of the learner in interpreting as well as recording information and in monitoring his or her own knowledge construction strategies;
- depends in part on the learner's ability to place an idea or skill within the framework of what is already known;
- occurs in context. In real life learning, tools, props, and other aids are often a part of the context in which one learns;
- is not automatically transferred from one context to another. There are probably severe limitations on generalized learning abilities;
- occurs most successfully in a social context. Knowledge that is posed, exposed and tested with others is more likely to have real and lasting meaning for the learner;
- proceeds from the general to the particular. Mastery of detail should not be expected before the big picture is explored.